Time and setting dependent instrument parameters and proofs of Bell-type inequalities

Karl Hess[1] and Walter Philipp[2]

[1] Beckman Institute, Department of Electrical Engineering and Department of Physics,University of Illinois, Urbana, Il 61801
[2] Beckman Institute, Department of Statistics and Department of Mathematics, University of Illinois,Urbana, Il 61801

Abstract

We show that all proofs of Bell-type inequalities, as discussed in Bell's well known book and as claimed to be relevant to Einstein-Podolsky-Rosen type experiments, come to a halt when Einstein-local time and setting dependent instrument parameters are included.

We have criticized in previous work [1]-[4] the various proofs of Bell-type inequalities as given in [5], [6]. Other authors have put forward related criticism of which we quote only some of the latest publications [7]-[9]. A very interesting discussion has developed during the last year [10]-[14]. In the present paper we concatenate our arguments into what we call "row" and "column" arguments. These arguments contain reasoning that is essential to any Bell-type proof. As we show, these arguments can not be completed when setting and time dependent instrument parameters are involved. This conclusion is obtained independently of our previous paper where we derive the quantum result [3]. We first review the parameter space introduced by Bell and our extension of this parameter space. We use a notation that is close to our previous papers [1]-[4]. However, for reasons of clarity we capitalize here all random variables and use the lower case for the values these random variables can assume.

Bell's [5] parameter random variables are essentially given by the functions $A_{\mathbf{a}}(\Lambda) = \pm 1, B_{\mathbf{b}}(\Lambda) = \pm 1$ that are related to the possible outcomes of spin measurements, with Λ being a parameter random variable that is related to information carried by the correlated particle pair that is sent out from a common source to two stations S_1 and S_2. We assume with Bell and others that the way Einstein-Podolsky-Rosen (EPR)- experiments are performed guarantees that Λ is independent of the instrument settings $\mathbf{a}, \mathbf{b}$. The form of the experiment was proposed originally by Bohm and Hiley [15].

We extend this parameter space [1]-[4] by adding setting and time dependent instrument parameter random variables $\Lambda^*_{\mathbf{a},t}$ specific to station S_1 and $\Lambda^{**}_{\mathbf{b},t}$ to station S_2. These variables may be stochastically independent of Λ. As an illustration, these variables $\Lambda^*_{\mathbf{a},t}$, $\Lambda^{**}_{\mathbf{b},t}$ can be thought of being generated by two computers with equal internal computer clock time but otherwise entirely independent. The variables could be represented by any programs that evaluate the input of $\mathbf{a}, t$ etc. We do not claim knowledge of any mathematical properties of these parameters as dictated by physics nor do we claim that they must exist in nature. We can currently not simulate the EPR experiment on such computers and never have claimed that we can. However, we postulate that any proof of Bell-type inequalities that claims relevance to locality questions must pass the test to be able to include such parameters. These parameters do obey Einstein locality and, therefore, must be covered by any EPR model that is constructed like Bell's and has the same purpose.

Such parameters require a setting and time dependent joint probability distribution

$$\rho_s(\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}, \Lambda) \tag{1}$$

The subscript s of ρ indicates the setting dependence. The setting and time dependent parameter random variables $\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}$ may be stochastically independent of Λ. Bell and all his followers exclude such parameters since they all have only one probability distribution $\rho(\Lambda)$ that does not depend on any setting. The instrument parameters that they do include are assumed to be conditionally independent given Λ i.e. if instrument parameters are included, then they have a product distribution (see [5] pp36):

$$\rho_s(\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}|\Lambda) = \rho_1(\Lambda^*_{\mathbf{a},t}|\Lambda) \cdot \rho_2(\Lambda^{**}_{\mathbf{b},t}|\Lambda) \tag{2}$$

Because we consider time correlations, Eq.(2) cannot hold for our setting and time dependent instrument parameters. Thus Bell-type proofs exclude a large set of joint probability distributions from their considerations. In the remainder of the paper, we deal with the question whether Bell-type proofs can be reformulated to include such setting and time dependent random variables. The answer will be in the negative.

Why were such instrument variables not considered? In our opinion because there was the belief that the requirement

$$A_{\mathbf{a}} = -B_{\mathbf{a}} \tag{3}$$

could not be fulfilled if correlations other than those given by Λ would be invoked. In fact, Bell himself writes on p 38 of [5] that for the case that Eq.(3) holds, "the possibility of the results depending on hidden variables in the instruments can be excluded from the beginning." Bell clearly did not consider the possibility of time correlations such as

$$A_{\mathbf{a}}(\Lambda, \Lambda^*_{\mathbf{a},t}, t_{meas.}) = -B_{\mathbf{a}}(\Lambda, \Lambda^{**}_{\mathbf{a},t}, t_{meas.}) \tag{4}$$

where the measurement times $t_{meas.}$ in the two stations are, for the same correlated pair, either the same (as indicated above) or at least linearly related and therefore can lead to Eq.(4) which is equivalent to Eq.(3) .

Can Bell-type proofs be saved by some reasoning in the extended parameter space? We show below, that all Bell type proofs contain what we call "row" and "column" arguments that can not be completed for the extended parameter space. We first consider a prototype of Bell's original proof arranged according to two types of reasoning (row and column):

(b1) The row argument of Bell:

For $x, y, z = \pm 1$ we have

$$|xz - yz| = |x - y| = 1 - xy \tag{5}$$

Substituting $x = A_{\mathbf{b}}(\Lambda)$, $y = A_{\mathbf{c}}(\Lambda)$ and $z = A_{\mathbf{a}}(\Lambda)$ gives

$$|A_{\mathbf{a}}(\Lambda)A_{\mathbf{b}}(\Lambda) - A_{\mathbf{a}}(\Lambda)A_{\mathbf{c}}(\Lambda)| = 1 - A_{\mathbf{b}}(\Lambda)A_{\mathbf{c}}(\Lambda) \tag{6}$$

(b2) The column argument of Bell: using the inequality

$$|\int f| \leq \int |f| \tag{7}$$

and the assumption that ρ is a probability density independent of the settings, we obtain

$$|\int (A_{\mathbf{a}}(\Lambda)A_{\mathbf{b}}(\Lambda) - A_{\mathbf{a}}(\Lambda)A_{\mathbf{c}}(\Lambda))\rho(\Lambda)d\Lambda| \leq 1 - \int A_{\mathbf{b}}(\Lambda)A_{\mathbf{c}}(\Lambda)\rho(\Lambda)d\Lambda \tag{8}$$

In view of the definition of the expectation value for the spin pair correlation $E(A_{\mathbf{a}} \cdot B_{\mathbf{b}})$, this yields Bell's inequality.

Does Bell's proof go forward when setting and time dependent instrument parameter random variables $\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}$ are included?

(hp1) The row argument with setting and time dependent instrument parameters leads to:

$$|A_{\mathbf{a},t_1}(...)A_{\mathbf{b},t_1}(...) - A_{\mathbf{a},t_2}(...)A_{\mathbf{c},t_2}(...)| =? \tag{9}$$

one can see immediately that inclusion of a time index that is related to the actual time of the measurement does not permit completion of Bell's reasoning as used in Eq.(6).

(hp2) The column argument with setting and time dependent instrument parameters:

Eq.(8) is based on the row argument expressed by Eq.(6). Because Eq.(6) can not be derived when our instrument parameters are included, the column argument can not be completed as well. In addition, the integration performed in Eqs.(7) and (8) involves now three different joint probability distributions and corresponding probability measures μ. Consequently, a fortiori, the integration does not lead to the inequality expressed in Eq.(8) and to Bell's inequality as can be seen by considering the following integrals on the left side of Eqs.(8):

$$|\int A_{\mathbf{a},t_1}(...)A_{\mathbf{b},t_1}(...)d\mu(\Lambda, \Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}) - \int A_{\mathbf{a},t_2}(...)A_{\mathbf{c},t_2}(...)d\mu(\Lambda, \Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{c}})| =? \tag{10}$$

Although this shows already clearly why Bell-type proofs do not go forward with time and setting dependent instrument parameters, we will repeat our reasoning more extensively by discussing tables of possible outcomes for the random variables which is a frequently used argument in proofs of Bell's inequality. We start again using only the parameter space of Bell.

(bt1) The row and column argument of Bell for tables of possible outcomes.

It is common practice [6] to form and discuss tables of possible outcomes that, invariably, involve in each row a certain sum of terms that we denote by Δ:

$$\Delta = A_{\mathbf{a}}(\Lambda)B_{\mathbf{c}}(\Lambda) - A_{\mathbf{a}}(\Lambda)B_{\mathbf{b}}(\Lambda) - A_{\mathbf{d}}(\Lambda)B_{\mathbf{b}}(\Lambda) - A_{\mathbf{d}}(\Lambda)B_{\mathbf{c}}(\Lambda) \tag{11}$$

At this point the following statistical argument is usually invoked in one form or another. If one considers possible outcomes λ^i that the random variable Λ may assume, then by the strong law of large numbers the values λ^i will appear approximately the same number of times for each of the setting pairs since these occur with the same probability. The reason for this argument is that the source parameter Λ does not depend on settings. However, this argument works only if the cardinality of the set $\{\lambda^i\}$ of values that Λ can assume is much smaller than the number of experiments performed. If all this is fulfilled, then one can reorder the possible outcomes of Λ in a thought experiment such that one has rows of four terms with the same value λ^i for each element of any given row:

$$\Delta^i = A_{\mathbf{a}}(\lambda^i)B_{\mathbf{c}}(\lambda^i) - A_{\mathbf{a}}(\lambda^i)B_{\mathbf{b}}(\lambda^i) - A_{\mathbf{d}}(\lambda^i)B_{\mathbf{b}}(\lambda^i) - A_{\mathbf{d}}(\lambda^i)B_{\mathbf{c}}(\lambda^i) = \pm 2 \tag{12}$$

Note that this possibility of reordering makes it unnecessary, at least in principle, to involve counterfactual arguments i.e. arguments of what would have happened if a different setting were chosen. One simply argues statistically that Λ will assume the same values no matter what the setting is and therefore one can reorder to obtain the table shown immediately below. We will see, however, that no such reordering is obvious for the extended parameter space and that one does need counterfactual reasoning to proceed with Bell-type arguments in that extended space. We will also see that mere counterfactual reasoning that still might be admissible is not sufficient to complete the Bell-type proofs. Because the above equation is true for each row, no separate column argument is needed to derive Bell-type inequalities. One can write out the following table and state: within Bell's original assumptions, the following table is "sampled" by the experimental procedure of EPR experiments:

$$\begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^i \\ \vdots \\ \lambda^M \end{bmatrix} \begin{bmatrix} +A_{\mathbf{a}}^1 B_{\mathbf{c}}^1 & -A_{\mathbf{a}}^1 B_{\mathbf{b}}^1 & -A_{\mathbf{d}}^1 B_{\mathbf{b}}^1 & -A_{\mathbf{d}}^1 B_{\mathbf{c}}^1 \\ +A_{\mathbf{a}}^2 B_{\mathbf{c}}^2 & -A_{\mathbf{a}}^2 B_{\mathbf{b}}^2 & -A_{\mathbf{d}}^2 B_{\mathbf{b}}^2 & -A_{\mathbf{d}}^2 B_{\mathbf{c}}^2 \\ \vdots & \cdots & \cdots & \vdots \\ +A_{\mathbf{a}}^i B_{\mathbf{c}}^i & -A_{\mathbf{a}}^i B_{\mathbf{b}}^i & -A_{\mathbf{d}}^i B_{\mathbf{b}}^i & -A_{\mathbf{d}}^i B_{\mathbf{c}}^i \\ \vdots & \cdots & \cdots & \vdots \\ +A_{\mathbf{a}}^M B_{\mathbf{c}}^M & -A_{\mathbf{a}}^M B_{\mathbf{b}}^M & -A_{\mathbf{d}}^M B_{\mathbf{b}}^M & -A_{\mathbf{d}}^M B_{\mathbf{c}}^M \end{bmatrix} = \begin{bmatrix} \pm 2 \\ \pm 2 \\ \vdots \\ \pm 2 \\ \vdots \\ \pm 2 \end{bmatrix} \tag{13}$$

Although the measurements of the various terms can only be made in sequence, at the end of the day one would have accumulated this table and obtained it by reordering provided that the element of physical reality that corresponds to a given λ^i could be somehow made visible. Of course, there may be some terms left over as implied by the application of the law of large numbers. However, the number of such incomplete rows is negligible for large M.

With time and setting dependencies permitted we need to invoke the following table in an attempt to proceed with reasoning similar to the above:

$$\begin{bmatrix} \lambda^1 & t_1 \\ \lambda^2 & t_2 \\ \vdots & \vdots \\ \lambda^i & t_i \\ \vdots & \vdots \\ \lambda^M & t_M \end{bmatrix} \begin{bmatrix} +A_{\mathbf{a}}B^*_{\mathbf{c}} & -A_{\mathbf{a}}B_{\mathbf{b}} & -A_{\mathbf{d}}B_{\mathbf{b}} & -A_{\mathbf{d}}B_{\mathbf{c}} \\ +A_{\mathbf{a}}B_{\mathbf{c}} & -A_{\mathbf{a}}B_{\mathbf{b}} & -A_{\mathbf{d}}B^*_{\mathbf{b}} & -A_{\mathbf{d}}B_{\mathbf{c}} \\ \vdots & \cdots & \cdots & \vdots \\ +A_{\mathbf{a}}B_{\mathbf{c}} & -A_{\mathbf{a}}B^*_{\mathbf{b}} & -A_{\mathbf{d}}B_{\mathbf{b}} & -A_{\mathbf{d}}B_{\mathbf{c}} \\ \vdots & \cdots & \cdots & \vdots \\ +A_{\mathbf{a}}B_{\mathbf{c}} & -A_{\mathbf{a}}B_{\mathbf{b}} & -A_{\mathbf{d}}B_{\mathbf{b}} & -A_{\mathbf{d}}B^*_{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} ? \\ ? \\ \vdots \\ ? \\ \vdots \\ ? \end{bmatrix} \tag{14}$$

Here the asterisk for one term per row is used to indicate that only one pair of settings is possible at a given time. The question-mark ? on the right hand side replaces now the ± 2 for reasons given immediately by discussing row and column arguments. Before doing so we digress to clarify the meaning of elements of physical reality that are related to the hidden parameters.

The EPR argument postulates a reality for the parameters or information that come with the particles from the source. If instrument parameters $\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}$ are introduced, then the reality of what is measured or the information content of what is measured depends now on both the information from the source and the information from the instruments in a mixed way. We are not talking about the reality of the information of the source alone but about that reality as "seen" through the instruments. That mixed information depends on the actual macroscopic settings of the instruments $\mathbf{a}, \mathbf{b}$ etc.. As a consequence this mixed reality can not exist for different settings at the same time. Therefore, all arguments involving all of Table(14) are now counterfactual. It is important then to explain how results from counterfactual arguments can be compared to experiments. Furthermore, one needs to deal with the fact that only the fraction of Table(14) marked by asterisks does correspond to possible values that the random variables may assume. Only one value may be assumed in a row, any second value is impossible, much as a coin toss can give only head or tail but not both at a time. The quantity Δ, representing a row, is therefore not even a random variable as can be seen from the textbook definition (P. Halmos): "a random variable is a function attached to an experiment - once the experiment has been performed the value of the function is known." These facts invalidate all Bell-type proofs known to us as can immediately be seen from the following.

(hpt1) The row argument for tables of possible outcomes with time and setting dependent instrument parameters:

The pair of settings that are chosen can be thought of as picked by the throw of a tetrahedal die. Therefore at a given time only one pair can be picked. Adding or counting more of the possible (for themselves) outcomes of a row amounts to the same as adding or counting different possible outcomes for tosses of the tetrahedal die (or of coin tosses as mentioned above) at a given time. The result of such a procedure is, in general, bound to be incorrect as can be shown by numerous examples. Therefore, the row argument does not work and cannot be used to compute the row-sum of ± 2 in Table(14).

(hpt2) The column argument for tables of possible outcomes with time and setting dependent instrument parameters:

Bell-type proofs need to show that the sampling of Table (14) leads to an expectation value of Δ that represents essentially Bell-type inequalities. However, Δ is not what is necessarily sampled by any procedure commensurate with the experiments as can be seen from the fact that one has four different integrals for the four expectation values of the spin pair correlation corresponding to the four columns e.g.

$$E(A_{\mathbf{a}}\cdot B_{\mathbf{b}}) = \int A_{\mathbf{a}}B_{\mathbf{b}}\rho_s(\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}, \Lambda)d\Lambda^*_{\mathbf{a},t}d\Lambda^{**}_{\mathbf{b},t}d\Lambda$$

or

$$E(A_{\mathbf{a}}\cdot B_{\mathbf{c}}) = \int A_{\mathbf{a}}B_{\mathbf{c}}\rho_s(\Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{c},t}, \Lambda)d\Lambda^*_{\mathbf{a},t}d\Lambda^{**}_{\mathbf{c},t}d\Lambda$$

and similar with other settings that appear in the Bell-type inequalities. Only if the joint distributions ρ_s are independent of the setting can the Bell-type inequalities be obtained the way they are usually proven.

We have thus excluded the main arguments that can be used to prove the Bell inequalities when time and setting dependent parameter random variables are involved i.e. the row and column arguments and the argument involving reordering. Below we attempt to shine some additional light on the failure of these arguments using slightly different viewpoints. It is important to be clear about the following fact. When time dependencies are included, the Bell-type arguments become necessarily counterfactual i.e. they involve considerations that can not be proven by experiment. This follows simply from the fact that any reasoning that involves considerations of different settings at the same time does not correspond to an actual experiment because at a given time only one setting of a macroscopic instrument is possible. The discussions above show also clearly that the technical expression “counterfactual” is, by itself, not sufficient to describe the extent of its meaning for a proof involving mathematical logic and deductions thereof. It definitely is still permissible to consider, as Einstein did, the possibility that a different setting could have been used. Recall Einstein’s original argument: if setting $\mathbf{a}$ is used in station S_1 and $A_{\mathbf{a}} = +1$ then one can predict with probability one the result $B_{\mathbf{a}} = -1$ in station S_2. Had instead setting $\mathbf{b}$ been used in S_1 and the result was $A_{\mathbf{b}} = +1$, then one can predict with probability one the result $B_{\mathbf{b}} = -1$ in S_2. We can not see anything wrong in the reasoning of this argument. However, if anyone starts adding or subtracting outcomes for settings that could have been chosen, then the result of such addition may not have any meaning and can be completely nonsensical. A simple example makes this clear. Consider ordering a meal from the menu in a restaurant. After the meal, the owner of the restaurant adds possible other choices that you could have made and presents the bill for all choices. Note, however, that in the standard text-book of Peres [6] such arguments occur in abundance. For instance on p 167 of [6] the author reports on "people who ponder over a menu in a restaurant"-"with no apparent ill effects", while on p 163 of [6] the author concedes with regard to his Eq.(6.24) that not all tests can be performed simultaneously. That fact does not prevent the author from totaling up the "results" of such non-performed experiments in his Eqs.(6.24) and (6.28). We therefore should name such situations as described above not just counterfactual but, rather "counter-syntaxial" [16] indicating that the counterfactual argument is compounded by the use of mathematical operations such as counting, adding, subtracting, etc. in a way that violates well established protocol. While counterfactual reasoning still may permit to continue with mathematical logic, counter-syntaxial reasoning invalidates any proof. The row argument in Bell-type proofs is counter-syntaxial and therefore mathematically inadmissible. What remains, therefore, is the column argument. In other words, Bell-proof supporters must show that the counterfactual Table (14) is still sampled in its entirety with high statistical accuracy because of the fact that a large fraction of the columns is sampled. However, we have shown above that this argument does not lead to Bell-type inequalities when setting and time dependent instrument parameters are involved because these imply the existence of setting dependent joint probability distributions. We have seen, however, in many discussions with colleagues, that these facts are not recognized because of a circular argument that can be summarized as follows.

It is often simply being assumed that what is sampled (e.g. in a Monte Carlo integration sense) by the procedure of random choice of setting, time and Λ is just Table (14), even if time correlations are involved. Others assume that if one would reorder the measured results one would invariably end up with a table equivalent to the full Table (14) ignoring the fact that the measurements are all at different times and that the reordering procedure would require extensive mathematical proof. Since this Table (14) leads immediately (by trivial algebra) to the Bell inequalities, the Bell inequalities are then assumed to be proven. This is clearly a logical circle. The task is to prove that any sampling procedure commensurate with the experiment or any such reordering procedure indeed gives a table equivalent to Table (14). That this is not possible by general mathematical methods can be seen from the following facts. Any sampling procedure that is commensurate with the experiments samples functions $A_{\mathbf{a}}, B_{\mathbf{b}}$ on a space that includes the respective settings, Λ and $\Lambda^*_{\mathbf{a},t}$, $\Lambda^{**}_{\mathbf{b},t}$ at time of measurement $t_{meas.}$. One can say that one samples the product $A_{\mathbf{a}}(\Lambda, \Lambda^*_{\mathbf{a},t}, t_{meas.}) \cdot B_{\mathbf{b}}(\Lambda, \Lambda^{**}_{\mathbf{b},t}, t_{meas.})$ by choosing randomly values of $\mathbf{a}, \mathbf{b}, \Lambda, \Lambda^*_{\mathbf{a},t}, \Lambda^{**}_{\mathbf{b},t}$ and $t_{meas.}$. Involving reasonable conditions for the parameters, this is indeed true and the result relates directly to the expectation value $E(A_{\mathbf{a}} \cdot B_{\mathbf{b}})$ of the spin pair correlation. To prove the Bell inequalities, however, one wants to sample Δ and not just the $A.B$ products. However, Δ is not randomly sampled, because each value of Δ involves four $A \cdot B$ products taken at the same time and with the same Λ

for four different settings while only one term of any row is sampled by the measurement procedure. Think of the sampling as a Monte Carlo integration procedure. To sample Δ, such an integration scheme must include random choices in any given row including two, three or four elements at the same time with the same value that Λ may have assumed and for different settings. However, only one such element is sampled per row. This problem is compounded by the fact that each column contains different stochastic variables that are setting dependent and follow different joint distributions. From this one can see that Δ is not sampled. Nor can we see any procedure of reordering that would map any possible outcomes that correspond to actual measurements onto Table (14). This brings home the importance of the definition of a random variable and the importance to adhere to that definition when probability theory is used. According to this definition, Δ is not a random variable and Bell-type proofs are therefore mathematically questionable in general. They become invalid when time and setting dependent instrument parameters are involved.

One can now ask the question whether there can be any other proof of the Bell-type which can deal with setting and time dependent instrument parameters. We have attempted to answer this question in reference [3] and we point the reader to this reference. There we show, that for any random sequence of setting pairs (that appear in the Bell inequalities) there exists a distribution of local setting and time dependent instrument parameters that leads to the quantum result for the expectation value $E(A_{\mathbf{a}} \cdot B_{\mathbf{b}})$ of the spin pair correlation. This model shows therefore that one can construct joint probability distributions that do not lead to Bell type inequalities and that Δ is not necessarily sampled by the experimental procedure. Therefore, assuming the validity of the results given in [3] , no Bell-type proof can be given because such proofs always need to show that Δ is sampled by the experimental procedure if the parameters involved obey Einstein locality. The discussion between Einstein and Bohr would then be undecidable by a probability model of the type that Bell suggested.

Acknowledgement: The work was supported by the Office of Naval Research N00014-98-1-0604.

References

[1] K. Hess and W. Philipp, quant-ph/0103028, March 7 (2001)

[2] K. Hess and W. Philipp, Proc. of the Nat. Academy of Sci. (USA), Vol. 98, 14224-14227, (2001)

[3] K. Hess and W. Philipp, Proc. of the Nat. Academy of Sci. (USA), Vol. 98, 1427-14234 (2001)

[4] K. Hess and W. Philipp, Europhys. Lett. Vol. 57, 775-781 (2002)

[5] J.S. Bell, "Speakable and Unspeakable in Quantum Mechanics", Cambridge University Press (1993)

[6] Asher Peres, "Quantum Theory: Concepts and Methods", Kluwer Academic Publications", Dordrecht-Boston-London (1998)

[7] L. Accardi and M. Regoli, Proc. Int. Conf. on Foundations of Probability and Physics, Vaxjo, Sweden, A. Khrennikov ed., pp 1-28 (2000)

[8] G. Adenier, Proc. Int. Conf. on Foundations of Probability and Physics, Vaxjo, Sweden, A. Khrennikov ed., pp 29-36 (2000)

[9] L. Accardi and M. Regoli, Proc. Int. Conf. on Foundations of Probability and Physics, Vaxjo, Sweden, A. Khrennikov ed., pp 15-30 (2001)

[10] N. David Mermin, quant-ph/0206118, (2002)

[11] N. David Mermin, quant-ph/0207140, (2002)

[12] K. Hess and W. Philipp, quant-ph/0208086, (2002)

[13] R. D. Gill, G. Weihs, A. Zeilinger and M. Zukovsky, Proc. Nat. Academy of Sci. (USA), Vol. 99, 14632-14635 (2002)

[14] K. Hess and W. Philipp, quant-ph/0209057

[15] D. Bohm and B. J. Hiley, Nonlocality in Quantum Theory Understood in Terms of Einstein's Nonlinear Field Approach, Foundations of Physics 11,pp 529-546 (1981)

[16] W. F. Stout suggested the word counter-syntaxial.